

N/A

1643

PAGE: 1

# RAW SEQUENCE LISTING

## PATENT APPLICATION US/09/305,390A

 DATE: 09/15/1999  
 TIME: 12:15:58

Input Set: I305390A.RAW

 This Raw Listing contains the General Information  
 Section and up to first 5 pages.

ENTERED

RECEIVED

APR 10 2000

TECH CENTER 1600/2900

```

1  <110> APPLICANT: Yamamoto, Hiroaki
2  <120> TITLE OF INVENTION: METHOD FOR PRODUCING OPTICALLY ACTIVE 4-HALO-3-HYDROXYBUT
3  ACID ESTER
4  <130> FILE REFERENCE: 06501/030001
5  <140> CURRENT APPLICATION NUMBER: US/09/305,390A
6  <141> CURRENT FILING DATE: 1999-05-05
7  <150> EARLIER APPLICATION NUMBER: JP 1998-126507
8  <151> EARLIER FILING DATE: 1998-10-21
9  <150> EARLIER APPLICATION NUMBER: JP 1998-300178
10 <151> EARLIER FILING DATE: 1998-10-21
11 <160> NUMBER OF SEQ ID NOS: 18
12 <170> SOFTWARE: FastSEQ for Windows Version 3.0
13 <210> SEQ ID NO 1
14 <211> LENGTH: 244
15 <212> TYPE: PRT
16 <213> ORGANISM: Escherichia coli
17 <400> SEQUENCE: 1
18 Met Asn Phe Glu Gly Lys Ile Ala Leu Val Thr Gly Ala Ser Arg Gly
19 1 5 10 15
20 Ile Gly Arg Ala Ile Ala Glu Thr Leu Ala Ala Arg Gly Ala Lys Val
21 20 25 30
22 Ile Gly Thr Ala Thr Ser Glu Asn Gly Ala Gln Ala Ile Ser Asp Tyr
23 35 40 45
24 Leu Gly Ala Asn Gly Lys Gly Leu Met Leu Asn Val Thr Asp Pro Ala
25 50 55 60
26 Ser Ile Glu Ser Val Leu Glu Lys Ile Arg Ala Glu Phe Gly Glu Val
27 65 70 75 80
28 Asp Ile Leu Val Asn Asn Ala Gly Ile Thr Arg Asp Asn Leu Leu Met
29 85 90 95
30 Arg Met Lys Asp Glu Glu Trp Asn Asp Ile Ile Glu Thr Asn Leu Ser
31 100 105 110
32 Ser Val Phe Arg Leu Ser Lys Ala Val Met Arg Ala Met Met Lys Lys
33 115 120 125
34 Arg His Gly Arg Ile Ile Thr Ile Gly Ser Val Val Gly Thr Met Gly
35 130 135 140
36 Asn Gly Gly Gln Ala Asn Tyr Ala Ala Ala Lys Ala Gly Leu Ile Gly
37 145 150 155 160
38 Phe Ser Lys Ser Leu Ala Arg Glu Val Ala Ser Arg Gly Ile Thr Val
39 165 170 175
40 Asn Val Val Ala Pro Gly Phe Ile Glu Thr Asp Met Thr Arg Ala Leu
41 180 185 190
42 Ser Asp Asp Gln Arg Ala Gly Ile Leu Ala Gln Val Pro Ala Gly Arg
43 195 200 205
44 Leu Gly Gly Ala Gln Glu Ile Ala Asn Ala Val Ala Phe Leu Ala Ser

```

PAGE: 2

RAW SEQUENCE LISTING  
PATENT APPLICATION US/09/305,390A

DATE: 09/15/1999  
TIME: 12:15:58

Input Set: I305390A.RAW

```

45          210          215          220
46      Asp Glu Ala Ala Tyr Ile Thr Gly Glu Thr Leu His Val Asn Gly Gly
47          225          230          235          240
48      Met Tyr Met Val
49      <210> SEQ ID NO 2
50      <211> LENGTH: 735
51      <212> TYPE: DNA
52      <213> ORGANISM: Escherichia coli
53      <220> FEATURE:
54      <221> NAME/KEY: CDS
55      <222> LOCATION: (1)...(732)
56      <400> SEQUENCE: 2
57      atg aat ttt gaa gga aaa atc gca ctg gta acc ggt gca agc cgc gga      48
58      Met Asn Phe Glu Gly Lys Ile Ala Leu Val Thr Gly Ala Ser Arg Gly
59          1          5          10          15
60      att ggc cgc gca att gct gaa acg ctc gca gcc cgt ggc gcg aaa gtt      96
61      Ile Gly Arg Ala Ile Ala Glu Thr Leu Ala Ala Arg Gly Ala Lys Val
62          20          25          30
63      att ggc act gcg acc agt gaa aat ggc gct cag gcg atc agt gat tat      144
64      Ile Gly Thr Ala Thr Ser Glu Asn Gly Ala Gln Ala Ile Ser Asp Tyr
65          35          40          45
66      tta ggt gcc aac ggc aaa ggt ctg atg ttg aat gtg acc gac ccg gca      192
67      Leu Gly Ala Asn Gly Lys Gly Leu Met Leu Asn Val Thr Asp Pro Ala
68          50          55          60
69      tct atc gaa tct gtt ctg gaa aaa att cgc gca gaa ttt ggt gaa gtg      240
70      Ser Ile Glu Ser Val Leu Glu Lys Ile Arg Ala Glu Phe Gly Glu Val
71          65          70          75          80
72      gat atc ctg gtc aat aat gcc ggt atc act cgt gat aac ctg tta atg      288
73      Asp Ile Leu Val Asn Asn Ala Gly Ile Thr Arg Asp Asn Leu Leu Met
74          85          90          95
75      cga atg aaa gat gaa gag tgg aac gat att atc gaa acc aac ctt tca      336
76      Arg Met Lys Asp Glu Glu Trp Asn Asp Ile Ile Glu Thr Asn Leu Ser
77          100          105          110
78      tct gtt ttc cgt ctg tca aaa gcg gta atg cgc gct atg atg aaa aag      384
79      Ser Val Phe Arg Leu Ser Lys Ala Val Met Arg Ala Met Met Lys Lys
80          115          120          125
81      cgt cat ggt cgt att atc act atc ggt tct gtg gtt ggt acc atg gga      432
82      Arg His Gly Arg Ile Ile Thr Ile Gly Ser Val Val Gly Thr Met Gly
83          130          135          140
84      aat ggc ggt cag gcc aac tac gct gcg gcg aaa gcg ggc ttg atc ggc      480
85      Asn Gly Gly Gln Ala Asn Tyr Ala Ala Ala Lys Ala Gly Leu Ile Gly
86          145          150          155          160
87      ttc agt aaa tca ctg gcg cgc gaa gtt gcg tca cgc ggt att act gta      528
88      Phe Ser Lys Ser Leu Ala Arg Glu Val Ala Ser Arg Gly Ile Thr Val
89          165          170          175
90      aac gtt gtt gct ccg ggc ttt att gaa acg gac atg aca cgt gcg ctg      576
91      Asn Val Val Ala Pro Gly Phe Ile Glu Thr Asp Met Thr Arg Ala Leu
92          180          185          190
93      agc gat gac cag cgt gcg ggt atc ctg gcg cag gtt cct gcg ggt cgc      624
94      Ser Asp Asp Gln Arg Ala Gly Ile Leu Ala Gln Val Pro Ala Gly Arg

```

RECEIVED

APR 10 2000

TECH CENTER 1600/2900

PAGE: 3

# RAW SEQUENCE LISTING PATENT APPLICATION US/09/305,390A

DATE: 09/15/1999  
TIME: 12:15:58

Input Set: I305390A.RAW

```

95          195          200          205
96      ctc ggc ggc gca cag gaa atc gcc aac gcg gtt gca ttc ctg gca tcc      672
97      Leu Gly Gly Ala Gln Glu Ile Ala Asn Ala Val Ala Phe Leu Ala Ser
98          210          215          220
99      gac gaa gca gct tac atc acg ggt gaa act ttg cat gtg aac ggc ggg      720
100     Asp Glu Ala Ala Tyr Ile Thr Gly Glu Thr Leu His Val Asn Gly Gly
101     225          230          235          240
102     atg tac atg gtc tga      735
103     Met Tyr Met Val
104 <210> SEQ ID NO 3
105 <211> LENGTH: 36
106 <212> TYPE: DNA
107 <213> ORGANISM: Artificial Sequence
108 <220> FEATURE:
109 <223> OTHER INFORMATION: Artificially synthesized primer sequence
110 <400> SEQUENCE: 3
111     aaaggatcca acaatgaatt ttgaaggaaa aatcgc      36
112 <210> SEQ ID NO 4
113 <211> LENGTH: 32
114 <212> TYPE: DNA
115 <213> ORGANISM: Artificial Sequence
116 <220> FEATURE:
117 <223> OTHER INFORMATION: Artificially synthesized primer sequence
118 <400> SEQUENCE: 4
119     tgcctcgagt tatcagacca tgtacatccc gc      32
120 <210> SEQ ID NO 5
121 <211> LENGTH: 248
122 <212> TYPE: PRT
123 <213> ORGANISM: Bacillus subtilis
124 <400> SEQUENCE: 5
125     Met Asp Met Leu Asn Asp Lys Thr Ala Ile Val Thr Gly Ala Ser Arg
126     1          5          10          15
127     Gly Ile Gly Arg Ser Ile Ala Leu Ala Leu Ala Lys Ser Gly Ala Asn
128     20          25          30
129     Val Val Val Asn Tyr Ser Gly Asn Glu Ala Lys Ala Asn Glu Val Val
130     35          40          45
131     Asp Glu Ile Lys Ser Met Gly Arg Lys Ala Ile Ala Val Lys Ala Asp
132     50          55          60
133     Val Ser Asn Pro Glu Asp Val Gln Asn Met Ile Lys Glu Thr Leu Ser
134     65          70          75          80
135     Val Phe Ser Thr Ile Asp Ile Leu Val Asn Asn Ala Gly Ile Thr Arg
136     85          90          95
137     Asp Asn Leu Ile Met Arg Met Lys Glu Asp Glu Trp Asp Asp Val Ile
138     100          105          110
139     Asn Ile Asn Leu Lys Gly Val Phe Asn Cys Thr Lys Ala Val Thr Arg
140     115          120          125
141     Gln Met Met Lys Gln Arg Ser Gly Arg Ile Ile Asn Val Ser Ser Ile
142     130          135          140
143     Val Gly Val Ser Gly Asn Pro Gly Gln Ala Asn Tyr Val Ala Ala Lys
144     145          150          155          160

```

PAGE: 4

RAW SEQUENCE LISTING  
PATENT APPLICATION US/09/305,390A

DATE: 09/15/1999  
TIME: 12:15:58

Input Set: I305390A.RAW

145 Ala Gly Val Ile Gly Leu Thr Lys Ser Ser Ala Lys Glu Leu Ala Ser  
146 165 170 175  
147 Arg Asn Ile Thr Val Asn Ala Ile Ala Pro Gly Phe Ile Ser Thr Asp  
148 180 185 190  
149 Met Thr Asp Lys Leu Ala Lys Asp Val Gln Asp Glu Met Leu Lys Gln  
150 195 200 205  
151 Ile Pro Leu Ala Arg Phe Gly Glu Pro Ser Asp Val Ser Ser Val Val  
152 210 215 220  
153 Thr Phe Leu Ala Ser Glu Gly Ala Arg Tyr Met Thr Gly Gln Thr Leu  
154 225 230 235 240  
155 His Ile Asp Gly Gly Met Val Met  
156 245

157 &lt;210&gt; SEQ ID NO 6

158 &lt;211&gt; LENGTH: 747

159 &lt;212&gt; TYPE: DNA

160 &lt;213&gt; ORGANISM: Bacillus subtilis

161 &lt;220&gt; FEATURE:

162 &lt;221&gt; NAME/KEY: CDS

163 &lt;222&gt; LOCATION: (1)...(744)

164 &lt;400&gt; SEQUENCE: 6

165 atg gat atg ctt aat gat aaa acg gct att gtc act ggc gca tcc cgc 48  
166 Met Asp Met Leu Asn Asp Lys Thr Ala Ile Val Thr Gly Ala Ser Arg  
167 1 5 10 15  
168 gga atc ggc cgc tca atc gcc ctt gct ctg gca aaa agc gga gca aat 96  
169 Gly Ile Gly Arg Ser Ile Ala Leu Ala Lys Ser Gly Ala Asn  
170 20 25 30  
171 gtt gtc gtg aac tac tcc ggc aat gaa gcg aaa gca aat gaa gtg gta 144  
172 Val Val Val Asn Tyr Ser Gly Asn Glu Ala Lys Ala Asn Glu Val Val  
173 35 40 45  
174 gat gaa atc aaa tca atg ggc aga aaa gca att gct gta aaa gcg gat 192  
175 Asp Glu Ile Lys Ser Met Gly Arg Lys Ala Ile Ala Val Lys Ala Asp  
176 50 55 60  
177 gta tca aat ccc gaa gat gta caa aac atg ata aaa gaa aca ttg tct 240  
178 Val Ser Asn Pro Glu Asp Val Gln Asn Met Ile Lys Glu Thr Leu Ser  
179 65 70 75 80  
180 gtt ttt tct acg att gac att ctg gtt aat aat gcg gga att aca aga 288  
181 Val Phe Ser Thr Ile Asp Ile Leu Val Asn Asn Ala Gly Ile Thr Arg  
182 85 90 95  
183 gac aat ctc atc atg aga atg aaa gaa gac gaa tgg gat gac gtc att 336  
184 Asp Asn Leu Ile Met Arg Met Lys Glu Asp Glu Trp Asp Asp Val Ile  
185 100 105 110  
186 aac att aac ctg aag ggt gtt ttc aac tgc aca aaa gct gtt aca aga 384  
187 Asn Ile Asn Leu Lys Gly Val Phe Asn Cys Thr Lys Ala Val Thr Arg  
188 115 120 125  
189 caa atg atg aaa cag cgt tca ggc cgc att att aac gta tcg tct atc 432  
190 Gln Met Met Lys Gln Arg Ser Gly Arg Ile Ile Asn Val Ser Ser Ile  
191 130 135 140  
192 gtc ggc gtc agc gga aac cct gga caa gcc aac tac gtg gct gca aaa 480  
193 Val Gly Val Ser Gly Asn Pro Gly Gln Ala Asn Tyr Val Ala Ala Lys  
194 145 150 155 160

PAGE: 5

# RAW SEQUENCE LISTING PATENT APPLICATION US/09/305,390A

DATE: 09/15/1999  
TIME: 12:15:58

Input Set: I305390A.RAW

```

195      gcc ggc gtc atc ggt tta acc aaa tct tct gct aaa gag ctc gcc agc      528
196      Ala Gly Val Ile Gly Leu Thr Lys Ser Ser Ala Lys Glu Leu Ala Ser
197              165              170              175
198      cga aat att acg gta aac gca ata gcg cca gga ttt atc tca act gat      576
199      Arg Asn Ile Thr Val Asn Ala Ile Ala Pro Gly Phe Ile Ser Thr Asp
200              180              185              190
201      atg aca gat aaa ctt gca aaa gac gtt caa gac gaa atg ctg aaa caa      624
202      Met Thr Asp Lys Leu Ala Lys Asp Val Gln Asp Glu Met Leu Lys Gln
203              195              200              205
204      att ccg ctc gcg cgc ttt ggt gaa cct agc gat gtc agc agt gtt gtc      672
205      Ile Pro Leu Ala Arg Phe Gly Glu Pro Ser Asp Val Ser Ser Val Val
206              210              215              220
207      acg ttc cta gct tca gag gga gct cgt tat atg aca ggc caa acg ctt      720
208      Thr Phe Leu Ala Ser Glu Gly Ala Arg Tyr Met Thr Gly Gln Thr Leu
209              225              230              235              240
210      cat att gac ggc gga atg gtg atg taa      747
211      His Ile Asp Gly Gly Met Val Met
212              245
213      <210> SEQ ID NO 7
214      <211> LENGTH: 33
215      <212> TYPE: DNA
216      <213> ORGANISM: Artificial Sequence
217      <220> FEATURE:
218      <223> OTHER INFORMATION: Artificially synthesized primer sequence
219      <400> SEQUENCE: 7      33
220      ggaccatgga tatgcttaat gataaaacgg cta
221      <210> SEQ ID NO 8
222      <211> LENGTH: 34
223      <212> TYPE: DNA
224      <213> ORGANISM: Artificial Sequence
225      <220> FEATURE:
226      <223> OTHER INFORMATION: Artificially synthesized primer sequence
227      <400> SEQUENCE: 8      34
228      gagaagcttc tcgagttaca tcaccattcc gccg
229      <210> SEQ ID NO 9
230      <211> LENGTH: 246
231      <212> TYPE: PRT
232      <213> ORGANISM: Ralstonia eutropha
233      <400> SEQUENCE: 9
234      Met Thr Gln Arg Ile Ala Tyr Val Thr Gly Gly Met Gly Gly Ile Gly
235              1              5              10              15
236      Thr Ala Ile Cys Gln Arg Leu Ala Lys Asp Gly Phe Arg Val Val Ala
237              20              25              30
238      Gly Cys Gly Pro Asn Ser Pro Arg Arg Glu Lys Trp Leu Glu Gln Gln
239              35              40              45
240      Lys Ala Leu Gly Phe Asp Phe Ile Ala Ser Glu Gly Asn Val Ala Asp
241              50              55              60
242      Trp Asp Ser Thr Lys Thr Ala Phe Asp Lys Val Lys Ser Glu Val Gly
243              65              70              75              80
244      Glu Val Asp Val Leu Ile Asn Asn Ala Gly Ile Thr Arg Asp Val Val

```

PAGE: 6

VERIFICATION SUMMARY  
PATENT APPLICATION US/09/305,390A

DATE: 09/15/1999

TIME: 12:15:58

Input Set: I305390A.RAW

Line ? Error/Warning

Original Text

-----